

Development of a Convolutional Neural Network for Classification of Type of Vessels

Ariel Victor do Nascimento¹, Marcus Pinto da Costa da Rocha¹, Valcir João da Cunha Farias², Miércio Cardoso de Alcântara Neto¹

¹ PPGENAV – Institute of Technology, Federal University of Para

² PGME – Institute of Exact and Natural Sciences, Federal University of Para

Received: 16 Aug 2022,

Receive in revised form: 18 Sep 2023,

Accepted: 20 Oct 2022,

Available online: 31 Jan 2023

©2023 The Author(s). Published by AI
Publication. This is an open access article
under the CC BY license

(<https://creativecommons.org/licenses/by/4.0/>).

Keywords— Convolutional Neural Network,
Artificial Intelligence; Deep Learning, Vessels

Abstract— In this paper, was used a method which use concepts of intelligence artificial, machine learning, deep learning for Classification of Type of Vessels. With a technique from deep learning called Convolutional Neural Network (CNN) was applied to recognize images to identify the type of ship and to use the same method to identify if the vessel. The CNN projected with determined 5 layers, the first layer containing 32 neurons, the second layer with 64 neurons, the third layer with 128 neurons, the fourth layer with 512 neurons. Activation functions for these specified layers contain the ReLU function. The fifth and last layer is the output layer is the output layer, so the number of neurons is equal to the number of vessel type. In our study six classes were used, which are the vessel types, in this layer the activation function was the Softmax. The CNN generate satisfactory results, where could get results of prediction with all corrects answers to identify the ship.

I. INTRODUCTION

The deep learning techniques will be used, which are called Convolutional Neural Networks (CNN). This programming technique is embedded within the Artificial Intelligence (AI) environment, which has been shown to be quite effective for image problems such as: Classification, pattern recognition [13], character extraction (OCR's) [11] and object detection [7], ship recognition using self-organizing distributed maps [9].

Based on the data provided by the Navy and the advancement of deep learning techniques, the objective is to develop a CNN code capable of classifying images of different type of vessels.

II. THEORETICAL REFERENCE

2.1. Types of Vessels according Normam

According Normas de Autoridade Marítima (NORMAM), there are many types of vessels, but in this

paper, was used six of them. The types chosen was: canoe, Catamaran, Ferry Boat, Yatch, Military Ship, Sailboat. In Figure 1 shows a sailboat called “Cisne Branco”, one example of the vessels.



Fig.1. Sailboat Cisne Branco

2.2. Intelligence Artificial, Machine Learning and Deep Learning

Siegel et. al. [14] explain the Artificial Intelligence is one of areas of computational science which focus on building machines and computers capable the simulate a type of intelligent behavior, making possible accomplish similar tasks the humans, like a recognizing voice, take decisions, visual perception, translating of languages. According of author Menezes [10], there are many areas where apply Artificial Intelligence.

For Bishop [4] and Almeida [1], Machine Learning is a type of system capable to get and accumulate knowledge and still yet improve the performance in specifics tasks. Are programs developed to learn to predict without have any supervision of programmer. This type of technique is applied in many areas, like processing images, recommendation songs, fraud detection.

Deep Learning (DL) built over neural network, being one type of machine learning, which is similar of human neurons. In Figure 2 show the artificial intelligence with subgroups [10].

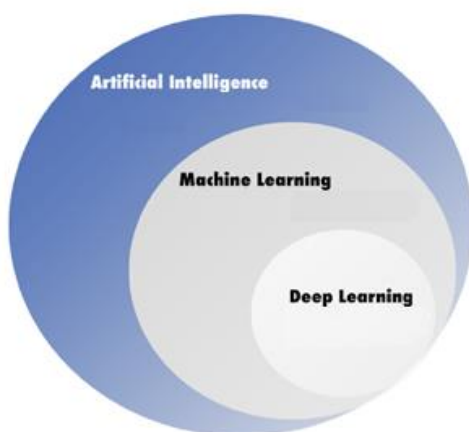


Fig.2. Subgroups of Artificial Intelligence

Anjos [2] declare that the deeper, or more levels, exist in the neural network, more is the quantity of operations and hence the need more computational power. The deep learning models are known too as neural networks because the first algorithms created to represent the biological learning, similar the brain.

2.3 Convolutional Neural Network

The great advantage of a convolutional neural network (CNN) on the other type of neural network it is your superior performance with the processing of image, speech, or audio signal inputs. The basic architecture consists of three types of layers, which are convolutional layer, pooling layer and fully connected layer.

Accord with IBM Cloud Education [8] The convolutional layer is the first layer of a CNN. While convolutional layers can be followed by additional convolutional layers or pooling layers, the fully connected layer is the final layer. With each layer, the CNN increases in its complexity, identifying greater portions of the image. Earlier layers focus on simple features, such as colors and edges. As the image data progresses through the layers of the CNN, it starts to recognize larger elements or shapes of the object until it finally identifies the intended object.

According to Almeida [1], one of the most important characteristics of CNN is the large number of architectures that can be configured, since networks have different layers that can be combined, including parameter variation.

For Bengio et. al. [3] the structure is composed of a series of consecutive layers convolution, pooling or variation, in the feature extraction phase. According to Anjos [2], the convolution layer the operation acts in 4 dimensions: filter width (x'), filter height (y'), filter channels (c) and number of filter (f):

$$F_{out} = \sum_{c=1}^C \sum_{x'=1}^{X'} \sum_{y'=1}^{Y'} Fin(c, x - x', y - y') W(c, x', y', f)$$

The output feature (F_{out}) of dimensions ($R_x \times R_y$) are produced by the convolution of the filters (W), of dimensions ($x' \times y'$), with the input feature (Fin), of dimensions ($X \times Y$).

III. METHODOLOGY

The methodology was separated in 5 steps. In step 1, images will be collected from many places, specific classes used on code, which are Military Ship, Ferry Boat, Canoe, Catamaran, Sailboat and Yacht, as seen previously. In step 2, a database will be created containing images from the internet, of the region.

All photographs are different for training and validation purposes, as it is important that the program can improve its parameters based on various points of view of the vessel type, without running the risk of training in a biased way.

After the elaboration of the database of images, proceeding to step 3, the development of the code for classification of images will be elaborated in the programming language python, which platform is considered open source, allowing users to use it free, in addition to of being able to use libraries, one of them being the TensorFlow (TF), created by the Google Brain team. TF is an open-source library for Machine Learning and there is an Application Programming Interface (API) that will be used for Deep Learning is Keras, which has a set of programming routines [6].

In this step, the process of the Convolutional Neural Network (CNN) will be discretized, including the number of neurons, layers, activation functions, Epoch and the form of displaying results on screen.

After setting the CNN and proceeding to step 4, the tool will run, where it will train and validate its parameters for classification, seeking to discard characteristics irrelevant features, for example the presence of water and prioritize specific contents, such as hull shape. For details about its accuracy, the percentage will be displayed on screen during all training periods.

Finally step 5, which consists of carrying out a test in which random images are presented, which are not included in the data base and generate a prediction of the results. In this case, if there is any non-conformity in the results, this error can lead to the interpretation that it may be related to the photographs in the dataset or the number of times generated.

This method will be applied both for the classification of the type of vessel and for the analysis of the vessel's freeboard.

IV. RESULT ANALYSIS

Images of six types of vessels were collected, as seen previously, and a database created for CNN to use for training.

For the method mentioned, in setup 3, the setting of the classification is specified categorically, because there are more than two classes to be evaluated. The images are very large dimensions, requiring a high demand of time and computational effort to process the pixels, they were resized to dimensions of 200 pixels long and 200 pixels wide, acceptable sizes for conventional computer. It was also determined that 5 layers of neurons would be used, the first containing 32 neurons, the second with 64 neurons, the third with 128 neurons, the fourth with 512 neurons. Activation functions for these specified layers contain the ReLU function ([15],[5]).

The fifth and last layer is the output layer is the output layer, so the number of neurons is equal to the number of inputs. In this case, 6 classes were used, which are the vessel types, with the Softmax activation function [12].

For the loss function, the Categorical Cross Entropy was used, which collaborated with evaluation of the epochs, to determine if the results were converging. After the training, in the last step, random images of vessels that are not in the database will be provided to CNN so that it can classify.



Fig.3. (a) War Ship; (b) Boat; (c) Yatch

The Convolutional Neural Network (CNN) was programmed to classify images according to the type of vessel. Currently the database consists of 120 photos, 60 for training and 60 for validation. Tests were performed with 30, 50 and 75 epochs. Was given to the CNN 6 photos in this order: War Ship (Figure 3-a), Boat (Figure 8-b), Yatch (Figure 3-c), Sailboat (Figure 9-a), Boat (Figure 9-b) and Catamaran (Figure 3-c).

At 30 epochs (Figure 5-a) the program get 4 of 5 correct answers, it had a mistake at last one picture. In this case, CNN answered Navio_Militar. At 50 epoch (Figure 5-b) was the same result at 30 epoch. For the last time, was executed with 75 epoch (Figure 5-c) which was showed the correct 5 of 5 answers.

It can be seen that there was an error in the last classification for training of 30 and 50 epochs. However, for 75 epochs, the network was correct in all its classifications



Fig.4. (a) Sailboat; (b) Boat; (c) Catamaran

V. CONCLUSION

Given the results presented, it can be seen that the neural network was successful in its first stage, which was to identify the types of vessels through images.

To increase the probability of correct predict in a low number of epochs it is necessary to increase the number of images in the database, since 120 photos is a very low number, and in some cases, there are databases with more than 5 thousand. From this, the database will be improved. Not only in relation to the number of existing photos, but also the increase in the types of vessels, in other words, CNN will be able to identify more classes.

In addition to identifying the type, it will also be enhanced with images of inclined vessels for the net to measure the freeboard.

With the expansion of the database, CNN will have an increase in accuracy in relation to its forecasts, being able to determine the type of stability risk in case the vessel is with low freeboard or inclined.

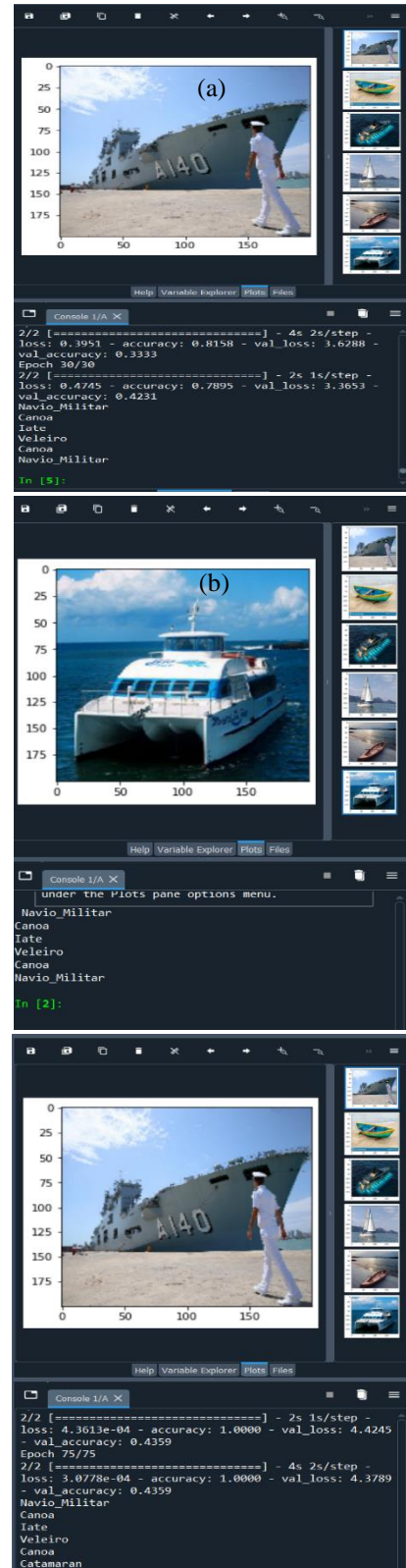


Fig.5. (a) 30 Epochs; (b) 50 Epochs; (c) 75 Epochs

ACKNOWLEDGEMENTS

The authors would like to thank the Postgraduate Program in Naval Engineering, Institute of Technology, Federal University of Pará.

REFERENCES

- [1] ALMEIDA, C. C. Identification and Classification of Images usando Convolutional Neural Network and Machine Learning: Implementation in na Embedded System. Thesis (PhD in Mechanical Engineering), State University of Campinas, Campinas, 2019.
- [2] ANJOS, J. R. L. InferenCNN: A Library for Inference of Multiplatform Convolutional Neural Network in OpenCL Dissertation (Master's degree in Computation of Science) Federal University of Pernambuco, Pernambuco, 2019.
- [3] BENGIO, Y. et al. Learning Deep Architectures for AI. Foundations and trends in Machine Learning, Now Publisher, Inc., v. 2, n. 1, p. 1- 127, 2009.
- [4] BISHOP, C. M. Pattern recognition and machine learning. Springer, 2006.
- [5] DAHL, G. E.; SAINATH, T. N.; HINTON, G. E. "Improving deep neural networks for LVCSR using rectified linear units and dropout" in International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013.
- [6] GOOGLE. An open source machine learning library for research and production. 2015. TensorFlow. Disponível em: <https://www.tensorflow.org/> . Acesso em: 08/07/2021
- [7] GOOGLE. Object Detection. 2017. Disponível em: https://github.com/tensorflow/models/tree/master/research/object_detection . Acesso em: 08/07/2021.
- [8] IBM Cloud Education Convolutional Neural Networks, 2020, https://www.ibm.com/cloud/learn/convolutional-neural-networks?sm_adopter=mh&mhsr=ibmsearch_a&mhq=Convolutional%20neural%20network
- [9] LOBO, V. S. Application of Self-Organizing Maps to the Maritime Environment. Information Fusion and Geographic Information Systems, Proceedings of the Fourth International Workshop, IF&GIS 2009, 17-20 May 2009.
- [10] MENEZES, C. Image Recognition by Convolutional Neural Networks (CNN) in Embedded Systems. 2018. Available in: <http://www.embarcados.com.br/webinars/webinar-reconhecimento-de-imagens-por-redes-neurais-convolucionais>. (accessed in 10th July, 2021)
- [11] MICROSOFT. Text Analytics. 2018. Disponível em: <https://azure.microsoft.com/pt-br/services/cognitive-services/text-analytics/>
- [12] NWANKPA, C. E. et al. Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. In: 2nd International Conference on Computational Sciences and Technology, pág. 124-133, Jamshoro, 2021.
- [13] Ronneberger, O., Fischer, P. and Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015. <https://arxiv.org/pdf/1505.04597.pdf>
- [14] SIEGEL, J. E. et al. Real-time deep neural networks for internet-enabled arc-fault detection. Engineering Applications of Artificial Intelligence, Elsevier, v. 74, 2018.
- [15] Zeiler, M. D.; Ranzato, M.; Monga, R.; Mao, M.; Yang, K.; Le, Q. V. e Hinton, G. E. "On rectified linear units for speech processing." Em International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013, pp. 3517-3521.